

DOCUMENT RESUME

ED 465 803

TM 034 190

AUTHOR Klein, Davina C. D.; Chung, Gregory K. W. K.; Osmundson, Ellen; Herl, Howard E.; O'Neil, Harold F., Jr.

TITLE Examining the Validity of Knowledge Mapping as a Measure of Elementary Students' Scientific Understanding. CSE Technical Report.

INSTITUTION California Univ., Los Angeles. Center for the Study of Evaluation.; National Center for Research on Evaluation, Standards, and Student Testing, Los Angeles, CA.

SPONS AGENCY Office of Educational Research and Improvement (ED), Washington, DC.

REPORT NO CSE-TR-557

PUB DATE 2002-04-00

NOTE 45p.

CONTRACT R305B60002

AVAILABLE FROM Center for the Study of Evaluation, National Center for Research on Evaluation, Standards, and Student Testing, Graduate School of Education & Information Studies, University of California at Los Angeles, 300 Charles E. Young Dr. North, Los Angeles, CA 90095-1522. Tel: 310-266-1532. For full text: <http://www.cse.ucla.edu/CRESST/Reports/TECH559.PDF>.

PUB TYPE Reports - Research (143)

EDRS PRICE MF01/PC02 Plus Postage.

DESCRIPTORS Comprehension; *Elementary School Students; Intermediate Grades; Multitrait Multimethod Techniques; *Reliability; *Scientific Principles; Scoring; Student Evaluation; *Validity

IDENTIFIERS *Knowledge Maps

ABSTRACT

Knowledge mapping is expected to measure deep conceptual understanding and allow students to characterize relationships among concepts in a domain visually. This research examined the validity of knowledge mapping as an assessment tool in science. The approach to investigating this validity was three-pronged. First, a model was outlined for the creation of knowledge mapping tasks, proposing a standard set of steps and using content area and educational experts to ensure the content validity of the measures. Next, a scoring method was developed to evaluate student performance. This report contains a discussion of the methods reliability and its relation to other possible scoring systems. Finally, the report presents the statistical results based on the participation of 56 fourth and fifth graders, including comparative analyses, the multitrait-multimethod (MTMM) validity analyses involving two traits (students understanding of hearing and of vision), and three different measurement methods (knowledge mapping, essay, and multiple-choice tasks), critical proposition analyses, and analyses of students, propositional elaborations. Results show knowledge maps to be sensitive to students competency level, with mixed MTMM results. The report concludes with a discussion of implications and directions for future work. (Contains 7 tables, 2 figures, and 33 references.) (Author/SLD)

CRESST

National Center for Research on Evaluation, Standards, and Student Testing

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

- ☒ This document has been reproduced as received from the person or organization originating it.
- ☐ Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND
DISSEMINATE THIS MATERIAL HAS
BEEN GRANTED BY

R. Dietel

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)

1



Examining the Validity of Knowledge Mapping as a Measure of Elementary Students' Scientific Understanding

CSE Technical Report No. 557

Davina C. D. Klein, Gregory K. W. K. Chung,
Ellen Osmundson, and Howard E. Herl
CRESST/University of California, Los Angeles

Harold F. O'Neil, Jr.
University of Southern California/CRESST



TM034190



UCLA Center for the Study of Evaluation

In Collaboration With:

UNIVERSITY OF COLORADO AT BOULDER • STANFORD UNIVERSITY • THE RAND CORPORATION
UNIVERSITY OF SOUTHERN CALIFORNIA • EDUCATIONAL TESTING SERVICE
UNIVERSITY OF PITTSBURGH • UNIVERSITY OF CAMBRIDGE



2

BEST COPY AVAILABLE

**Examining the Validity of Knowledge Mapping
as a Measure of Elementary Students'
Scientific Understanding**

CSE Technical Report No. 557

Davina C. D. Klein, Gregory K. W. K. Chung,
Ellen Osmundson, and Howard E. Herl
CRESST/University of California, Los Angeles

Harold F. O'Neil, Jr.
University of Southern California/CRESST

April 2002

Center for the Study of Evaluation
National Center for Research on Evaluation,
Standards, and Student Testing
Graduate School of Education & Information Studies
University of California, Los Angeles
Los Angeles, CA 90095-1522
(310) 206-1532

Project 1.3. Technology in Action
Eva Baker, CRESST/University of California, Los Angeles, Project Director

Copyright © 2002 The Regents of the University of California

The work reported herein was supported under the Educational Research and Development Centers Program, PR/Award Number R305B60002, as administered by the Office of Educational Research and Improvement, U. S. Department of Education.

The findings and opinions expressed in this report do not reflect the positions or policies of the National Institute on Student Achievement, Curriculum, and Assessment, the Office of Educational Research and Improvement, or the U. S. Department of Education.

THE VALIDITY OF KNOWLEDGE MAPPING AS A MEASURE OF ELEMENTARY STUDENTS' SCIENTIFIC UNDERSTANDING¹

Davina C. D. Klein, Gregory K. W. K. Chung,
Ellen Osmundson, and Howard E. Herl²
CRESST/University of California, Los Angeles

Harold F. O'Neil, Jr.
University of Southern California/CRESST

Abstract

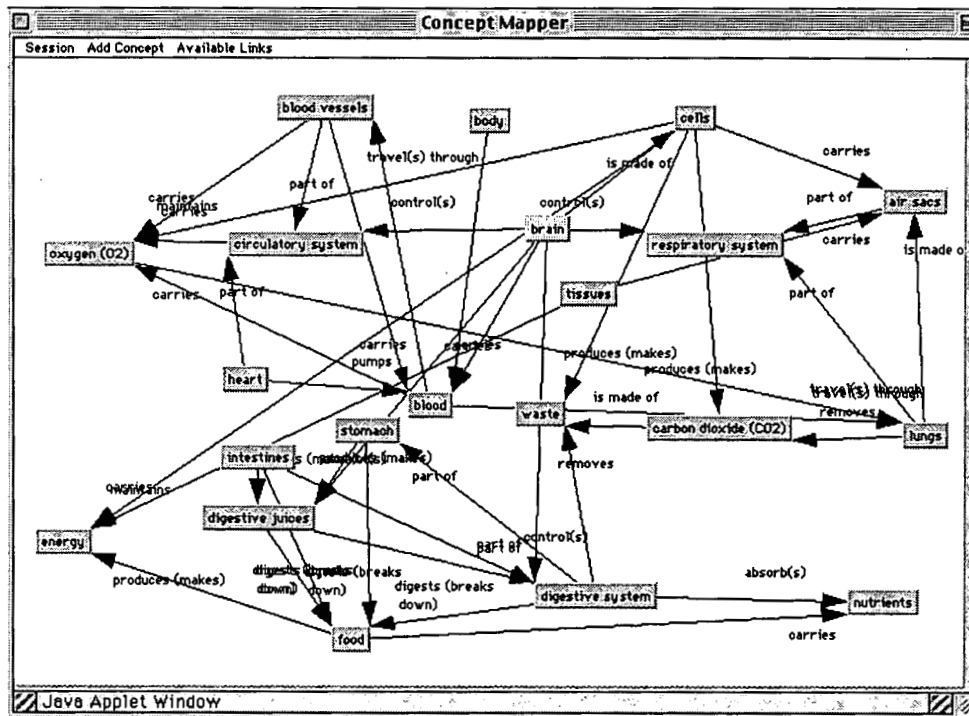
Although first popular as an instructional tool in the classroom, knowledge mapping has been used increasingly as an assessment tool. Knowledge mapping is expected to measure deep conceptual understanding and allow students to characterize relationships among concepts in a domain visually. Our research examines the validity of knowledge mapping as an assessment tool in science. Our approach to investigating this validity is three-pronged. First, we outline a model for the creation of knowledge mapping tasks, proposing a standard set of steps and using content area and educational experts to ensure the content validity of the measures. Next, we describe a scoring method used to evaluate student performance, including a discussion of the method's reliability and its relationship to other possible scoring systems. Finally, we present our statistical results including comparative analyses, our multitrait-multimethod validity analyses involving two traits (students' understanding of hearing and of vision) and three different measurement methods (knowledge mapping, essay, and multiple-choice tasks), critical proposition analyses, and analyses of students' propositional elaborations. Results show knowledge maps to be sensitive to students' competency level, with mixed MTMM results. We conclude with a discussion of implications and directions for future work.

As evidence mounts regarding the limitations of standardized multiple-choice testing, educators increasingly are looking for ways of assessing students' scientific conceptual understanding that may be missed by more traditional measures. Students' performance on knowledge mapping tasks has emerged as one possible

¹ We wish to acknowledge and thank Joanne Michiuye and Ali Abedi of CRESST/UCLA for their invaluable technical support, Robby Klein and Andrew Shpall for their content expertise, and Uyen Bui for her interest, participation, and content expertise. Finally, we are deeply grateful to Sharon Sutton, Jan Cohn, and their students for their assistance and participation in this research.

² Howard Herl is now at the Los Angeles County Office of Education.

source of information regarding their scientific knowledge. Knowledge mapping³ allows students to represent their understanding graphically, using nodes to represent main ideas and links to represent the relationships between the ideas. Students construct knowledge maps to demonstrate their knowledge of the important relationships among main ideas within a domain (see Figure 1 for a sample knowledge map).



³ We use the term *knowledge mapping* rather than *concept mapping* because we believe it to be a broader term, encompassing both conceptual knowledge and other types of knowledge (e.g., procedural knowledge). However, these tasks certainly could be characterized as concept-mapping tasks.

Research indicates that students who use knowledge maps are better at integrating, organizing, comprehending, retaining, and recalling new material (Armbruster & Anderson, 1984; Holley & Dansereau, 1984; Jonassen et al., 1993; Okebukola & Jegede, 1988). As instruction and assessment intersect, the use of knowledge mapping has moved into the assessment arena (Baker, Niemi, Novak, & Herl, 1992; Chung, O'Neil, & Herl, 1999; Herl, 1995; Herl, Baker, & Niemi, 1996; Jonassen et al., 1993; McClure, Sonak, & Suen, 1999; Novak, 1998; Ruiz-Primo & Shavelson, 1995; Ruiz-Primo, Shavelson, & Shultz, 1997). Researchers have suggested that as an assessment task, knowledge mapping can elicit students' deep conceptual understanding. Like an essay task, a knowledge map allows students to demonstrate their understanding of relationships between complex concepts. However, unlike essays, which are usually scored by human raters and often at considerable expense (Hardy, 1995, estimates the cost of scoring essays to range from \$3 to \$6 per essay, using a holistic rubric and a rater scoring rate of 12 minutes per essay), knowledge maps can be scored via computer.

The advantages of knowledge maps as assessment tools are numerous. The mechanics of constructing knowledge maps are easy, and students are quick to learn how to construct both paper-and-pencil and computer-based knowledge maps. In addition, having been used successfully as a learning device, knowledge maps allow a link between instruction and assessment, boosting the face validity of knowledge mapping as an assessment tool. Further, due to recently developed computerized scoring solutions, scoring of knowledge maps is straightforward and cost-effective. Finally, and most importantly, we believe knowledge mapping allows us to evaluate deep conceptual understanding. However, although research studies have documented the benefits of knowledge mapping in instructional settings, issues of validity in assessment settings remain unanswered.

Our hypothesis is that knowledge maps yield information about student competency that is different from, yet overlapping with, information revealed by essays and multiple-choice tasks. Although knowledge mapping may not be able to give us the kind of detailed and in-depth information about student understanding possible using an essay task, it seems nonetheless to allow us a glimpse into students' deep understanding in a domain—one that multiple-choice tasks may not allow.

Our approach to investigating the validity of knowledge mapping as an assessment task is three-pronged. First, we outline a model for the creation of

knowledge mapping tasks, proposing a standard set of steps and using content area and educational experts to ensure the content validity of the measures. Next, we describe the scoring method used to evaluate student performance, including a discussion of the method's reliability and its relationship to other possible scoring systems. Finally, we present our statistical results including comparative analyses, our multitrait-multimethod validity analyses, which involve two traits (students' understanding of hearing and of vision) and three different measurement methods (knowledge mapping, essay, and multiple-choice tasks), critical proposition analyses, and analyses of students' propositional elaborations. By addressing these three properties of knowledge maps—their creation, their scoring, and their relationship to other measures of understanding—we aim to gather evidence on the validity of knowledge maps for measuring students' scientific understanding.

Development Model for Knowledge Mapping Tasks

Our model for the creation of knowledge mapping tasks uses both content area experts (e.g., scientists, historians, mathematicians) and instructional experts (e.g., teachers, researchers) working together to devise a task to measure students' conceptual understandings. Underlying the mapping task is a scoring system that credits students for creating propositions that are like those of experts; students' maps are scored against multiple experts' maps rather than against one correct answer.

The development of knowledge mapping assessment tasks involves five steps (see Table 1 for process summary). First, the topic area to be assessed is specified. Experts in the field (e.g., medical practitioners, historians, scientists) are asked to generate lists of the most important “big ideas” within the topic area. Previous expert-novice research has shown that experts generally organize their knowledge around certain key principles or important ideas, then link these ideas together in a principled manner (Chase & Simon, 1973; Chi, Feltovich, & Glaser, 1981; Chi, Glaser, & Farr, 1988). Thus, these concepts serve to anchor the knowledge map within a particular topic area. Next, supporting classroom material, textbooks, and teacher interviews are used to tailor the experts' lists of concepts to the particular student audience. For example, in this study, medical students first described the hearing process (using terms such as *sound waves*, *ear drum*, and *pitch*); then teachers adjusted the material to the appropriate grade level (deleting, for instance, the term *pitch*).

Table 1
Development Model for Knowledge Mapping Task

Step	Process
1	Specify topic area and ask experts to generate important concepts.
2	Review and tailor concept list to particular student audience.
3	Construct preliminary knowledge maps with concept list in order to generate set of linking words.
4	Review and tailor link list to particular student audience.
5	Construct final knowledge maps to be used for scoring purposes.

After identifying an appropriate set of important ideas or concepts, experts are asked to create preliminary knowledge maps using their own linking words. Concepts are connected with directional lines to create concept-link-concept sets, termed *propositions*,⁴ which act to form a sentence. For instance, one might create the proposition *nerve sends message to brain*. Expert maps are compared, links are discussed with the teachers, and a final list of links appropriate to the students' level is generated from this information. For instance, in our hearing task, the list included links such as *is part of*, *vibrates*, and *is connected to*.

In the last step of the process, experts again construct knowledge maps in the topic area, this time using the specified set of final concepts and links. These expert maps are later used to score student maps.

Scoring of Knowledge Mapping Tasks

The method of scoring student outcome maps is a crucial concern when using knowledge mapping in an assessment setting. Unlike traditional multiple-choice tests, there is clearly no one "correct" knowledge map for a given topic domain. A variety of student knowledge maps containing different sets of propositions all could score well in comparison to experts' maps. On the other hand, scoring of knowledge maps also differs significantly from scoring of other performance-based assessment tasks, such as the essay tasks used in this study. Whereas raters can be trained to score explanation tasks holistically—using benchmark papers and an explicit scoring rubric—cost factors and the complexity of a 10-term or 15-term

⁴ We use the terms *proposition* and *link* at times interchangeably. Technically, a link is only the connection between two concept words; in practice, we use the term *link* to describe both the link between terms and the concept-link-concept proposition set. In context, it will be clear to which we refer.

knowledge map with numerous links preclude this kind of evaluation approach at any more than a rudimentary level (e.g., McClure et al., 1999). Regardless of the scoring scenario, the validity and reliability of knowledge map scores are of utmost importance.

Research on knowledge mapping has often focused on hierarchical knowledge maps, a more restrictive type of map (see, for instance, Markham, Mintzes, & Jones, 1994; Novak & Gowin, 1984; Wallace & Mintzes, 1990). Scoring methods for these types of maps generally include using the number of hierarchical and cross-links as measures of content knowledge. However, because much content knowledge need not be represented hierarchically (e.g., hearing and vision processes) and since less ordered methods of constructing knowledge maps allow for hierarchical structure, we do not find scoring schemes that expect hierarchically structured knowledge maps to be of much benefit in scoring less structured maps. Other research dealing with less structured knowledge maps (for example, Austin & Shore, 1995) has used more basic scoring techniques, often assigning scores based on the number of concepts and/or links, number of "good links," and so on—a somewhat limited and often arbitrary system. Finally, some researchers have used expert maps as the criteria for scoring students' maps (Herl, 1995; Herl et al., 1996; McClure et al., 1999; Ruiz-Primo, Schultz, & Shavelson 1997).

Using criterion maps has the advantage of allowing researchers to score student knowledge maps in various ways. For instance, rather than simply adding up terms and links, the degree to which a student's map matches an expert's map can vary by the definition of "match." Ruiz-Primo, Schultz, and Shavelson (1997) used three different map scores: (a) a total score, defined as the number of valid student links; (b) a congruence score, defined as the proportion of valid student links to all criterion links; and (c) a salience score, defined as the proportion of valid student links to all student links. Herl and his colleagues (1996) used a matching algorithm involving multiple expert maps in the scoring of each student map. Thus, matching entailed having the same proposition as any expert, and the degree of matching was weighted by the proportion of experts a student matched. Herl et al. calculated two related mapping scores: (a) a stringent semantic content score, based on exact link matches between student links and expert links; and (b) a categorized semantic content score, based on students matching some set of possible links (e.g., the causal set of links included *contributed to*, *encourages*, and *led to* links).

The matching-to-expert-performance approach also has other advantages. First, it uses content experts (either instead of, or in addition to, teachers or assessment developers) to establish accurate connections. This adds to the validity of the assessment score. Whereas teacher maps might contain misconceptions or less exact links, expert maps should include all the important, relevant interconnections between the given terms, without any errors or omissions. Furthermore, research has shown considerable differences between expert and novice content knowledge, with expert content knowledge exhibiting deeper, more connected, interrelated patterns (Chi et al., 1988; Gentner & Stevens, 1983; Glaser & Bassok, 1989). Second, by using many experts, this scoring approach allows for a more open-ended response—one in which students can show their competency in a variety of ways—while still ensuring that only student responses matching expert standards receive high marks. Finally, using this matching approach, knowledge map scores have been found to rank students consistently relative to one another as well as to provide stable individual student performance measures (McClure et al., 1999; Ruiz-Primo, Schultz, & Shavelson, 1997). Map scores also have been found to positively, moderately correlate both with essay tasks (Herl, 1995) and with multiple-choice tests (Ruiz-Primo, Schultz, & Shavelson, 1997), thereby providing some evidence for the validity of the scoring method.

In this study, knowledge maps were scored online using a basic, stringent match-to-expert algorithm.⁵ Maps were scored by comparing each student map to two expert maps. Students received half a point for each link that matched with an expert link, and a full point if the link matched with both experts' links. Students' total summed scores were used in the analyses.

In addition to this scoring method, we attempted to improve upon this approach by modifying the scoring scheme in two different but related ways. Our first deviation, which we term *supplementary credit*, gave students credit for (a) links that were more valued by experts ("critical links") and (b) links that were less like what experts might include. To this end, we asked experts to complete two tasks: (a) Experts selected the four to six most important propositions in their knowledge maps—those that best captured the key or crucial relationships in each of their knowledge maps, and (b) experts rated all other propositions that they did not

⁵ A categorized semantic score was ruled out because the links selected for our tasks were not conducive to set categorization; linking words in our science tasks were more specific than those used in previous studies.

include in their maps as either invalid (nonsensical) or valid (correct). Student map scores were then calculated by summing the proposition scores in each student's map; this coding approach assigned no points for an invalid proposition, one point for a valid proposition, two points for a normal proposition found in either expert's map, and three points for a proposition identified as critical in an expert's map. This approach eliminated two possible problems with our original scoring method. Students were given credit for knowing basic scientific facts—valid links that experts do not generally bother putting into their knowledge maps (for instance, *ear drum is part of ear*). Conversely, since experts identified key, critical links, students were given supplementary credit for identifying particularly crucial conceptual links—adding more information to our prior scoring approach.

Our second deviation from the stringent match-to-expert scoring approach, which we term the *proposition-quality rating system* (Osmundson, 1998; Osmundson, Chung, Herl, & Klein, 1999), asked an expert to consider each and every possible concept-link-concept proposition and assign it to one of four categories, describing the type of scientific understanding it revealed: illogical (type 0), pragmatic (type 1), scientific (type 2), or principled (type 3). Although this approach is similar to the first deviation described above, it looks at the knowledge maps from a different perspective, focusing not only on what should be included in a map but also on the level of abstraction and depth of the underlying content.

Method

Participants

Fifty-six 4th- and 5th-grade science students participated in this study. Students were of mixed ethnicity and socioeconomic status. Median family income was less than \$50,000; the range was less than \$20,000 per year to more than \$250,000 per year. All students attended a university laboratory school in Los Angeles.

Students were drawn from two intact classrooms of the same instructor. The classes were equivalent in terms of achievement and other academic indicators, according to the classroom teacher's informal assessment and standardized test scores. The mean SAT-9 total reading stanine scores were 6.20 ($SD = 1.50$, range = 2 to 9, $n = 25$; $F(1, 50) = 1.14$, $p = .29$) and 5.67 ($SD = 2.04$, range = 0 to 8, $n = 27$; $F(1, 50) = 0.04$, $p = .85$). The mean SAT-9 total math stanine scores were 6.28 ($SD = 1.57$, range = 3 to 9, $n = 25$) and 6.18 ($SD = 2.00$, range = 2 to 9, $n = 27$). Students were

familiar with computers, did not have prior knowledge mapping experience, and recently had finished learning about vision and hearing processes in their science classes.

Design

We used a 3 x 2 fully crossed design for this study. Each student completed six tasks, including three types of tasks (knowledge mapping, multiple choice, and essay) in two different topic areas (hearing and vision). To control for topic and task differences, task order was counterbalanced, as was the order of topic presentation.

Measures

We constructed three types of measures for this study: knowledge maps, essays, and a multiple-choice test. Each of these measures had two components, one addressing the hearing process and the other, the vision process.

Knowledge maps. We used our task creation model to generate two different science knowledge mapping tasks: a hearing task and a vision task. Two 4th-year medical students from a major university were recruited to serve as content area experts; the classroom teacher and the researchers served as instructional experts. The hearing task included 15 concepts and 12 links; the vision task consisted of 15 concepts and 14 links. All mapping tasks were administered online using the Knowledge Mapper software (described below). Table 2 presents the lists of concepts and links available to the students. Appendix A contains the student knowledge mapping prompts.

In conjunction with the knowledge mapping tasks, we created both online and offline worksheets to collect students' explanations, or "elaborations," of the links they considered most important in their knowledge maps (see Appendix B). These measures allowed us to investigate the kinds of thought processes underlying students' knowledge mapping performance.

Essays. Two essay tasks were designed in collaboration with the classroom teacher to measure students' understandings of the hearing and vision processes (see Appendix C for the essay task prompts). These assessments were based on prior work at CRESST using explanation tasks to assess deep conceptual understanding in a domain (e.g., Baker, Aschbacher, Niemi, & Sato, 1992; Baker & Niemi, 1998). The essay tasks were administered as paper-and-pencil measures.

Table 2

Hearing and Vision Concept and Link Lists

Hearing		Vision	
Concepts	Links	Concepts	Links
Anvil	Bend(s)	Brain	Bend/refract(s)
Brain	Control(s)	Color	Control(s)
Cochlea	Gather/receive(s)	Cones	Create(s)
Ear	Has to do with	Cornea	Focus(es)
Ear canal	Is connected (to)	Eye	Gather/receive(s)
Eardrum	Is filled with	Image	Has to do with
Fluid	Is made up of	Iris	Hit(s)
Hair	Is part of	Lens	Is connected (to)
Hammer	Protect(s)	Light	Is made up of
Inner ear	Send(s) messages to	Muscles	Is part of
Middle ear	Travel(s) through	Object	Move(s)
Nerve	Vibrate(s)	Optic nerve	Protect(s)
Outer ear		Pupil	Send(s) messages to
Sound waves		Retina	Travel(s) through
Stirrup		Rods	

Multiple-choice test. A multiple-choice test was created (also in collaboration with the classroom teacher) to measure students' basic understanding of the hearing and vision processes. The test consisted of 12 hearing items and 12 vision items and is included in Appendix D.

Knowledge Mapping System

Figure 2 shows the main user interface of our Web-based knowledge mapping system. The system has an easy-to-use, point-and-click interface requiring only a mouse. Students add concepts to their maps by selecting terms from a menu. Students link concepts together by dragging the mouse from one concept to the next, then clicking on the desired link from a pop-up menu (thus creating a concept-link-concept proposition such as *outer ear gathers sound waves*). Concepts are used only once, turning to gray on the concept menu to indicate their prior use; links may be reused and can be created between any two concepts. Student training on the use of the tool required about 10 minutes and included instruction both on the basics of knowledge mapping (e.g., What is the purpose of a knowledge map? What is a link? How do you construct a map?) and on how to use the system.

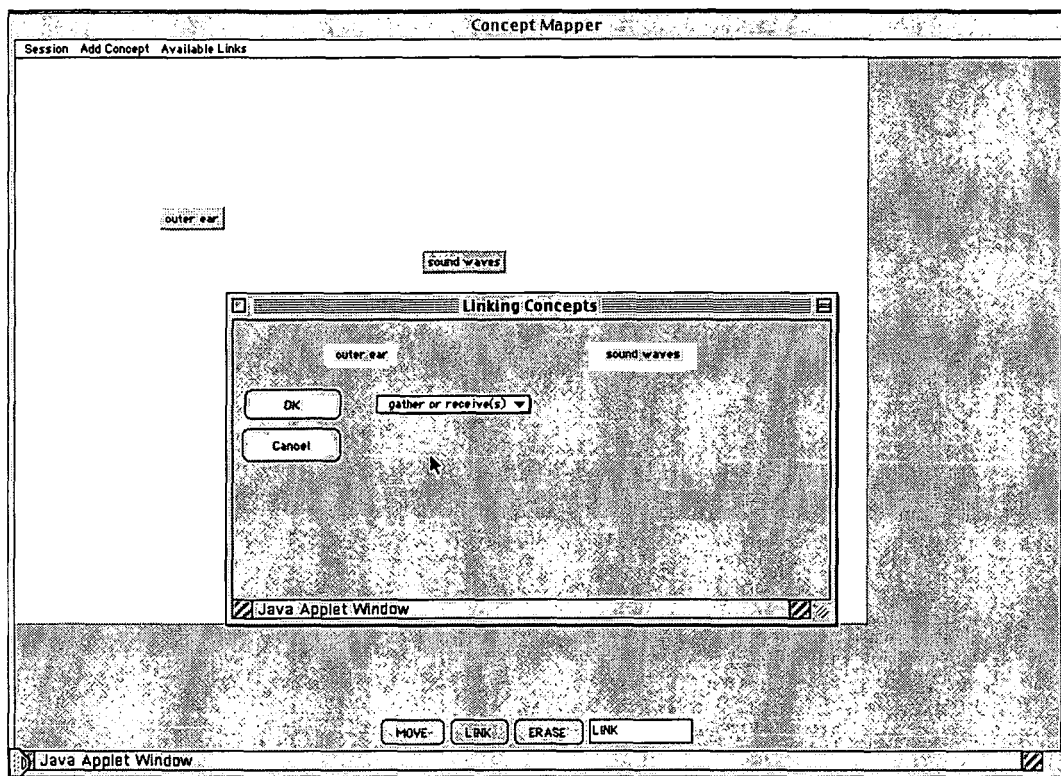


Figure 2. The Knowledge Mapper.

Procedures

Data were collected over the course of a week. Students had approximately 25 minutes to complete each task. Essays and multiple-choice tasks were completed using paper and pencil; knowledge mapping tasks were completed online using the Web-based knowledge mapping tool.

Knowledge mapping tasks. The mapping tasks were presented to students within a prompt context: Students were told to create a map to organize all of the important ideas about hearing (or vision) and to show how the different parts of the hearing (or vision) process relate to one another ("go together"). The prompt created a familiar context for the student to frame the task; students were asked to help a friend who had missed the last two months of school understand how our ears (or eyes) work.

Knowledge maps were scored via computer with a match-to-expert algorithm using two expert knowledge maps for each topic area as templates. Possible scores for the student maps ranged from 0 to 14.5 for hearing and 0 to 13.5 for vision. Inter-

expert agreement was calculated by examining the percentage of links (relative to the total number of links) that both experts included in their maps.⁶ Inter-expert agreement was found to be .72 for hearing and .77 for vision.

In addition to the match-to-expert scoring algorithm, we used both the supplementary credit and the proposition-quality rating systems in an attempt to further improve our prior scoring algorithm. Our results were somewhat surprising. We used Pearson correlation coefficients to test for the similarity between the various coding systems. We found the three coding schemes to be highly correlated, with Pearson coefficients ranging from .75 to .93 and a mean coefficient of .87 (see Table 3). Thus, although we expected to better our prior scoring system, we found that our original system may suffice. It is simpler to apply than the other two and supplies roughly the same information about student competency.

Upon completion of their knowledge maps, students were asked to review their maps and to select the two links they considered most important. They then wrote an explanation of each of these two links. Half of these elaborations were collected online (with students typing in their explanations) and half were written out on paper to control for students' typing abilities.

We examined the elaborations students composed to explain their links. We then coded the elaborations into five categories in order to better understand what information might be lost when using only a graphical representation of the link (rather than a labeled line plus a textual explanation). Each elaboration was coded

Table 3

Correlations Between Knowledge Mapping Scoring Systems (Hearing/Vision)

Scoring approaches	Original	Supplementary credit	Quality rating
Original	—	.92/.84	.86/.75
Supplementary credit	.92/.84	—	.93/.89
Quality rating	.86/.75	.93/.89	—

⁶ Note that the *is part of* and *is made up of* links are reciprocal; that is, one could create links that show *the ear is made up of* the inner, middle, and outer ears or that the inner, middle, and outer ears *are part of the ear*. Since they are synonymous, these reciprocal links were considered equivalent for purposes of calculating inter-expert agreement.

into one of five categories: (1) no new information/repetition of proposition, (2) elaboration and proposition both correct, (3) elaboration and proposition both incorrect, (4) correct elaboration paired with incorrect proposition, and (5) incorrect elaboration paired with correct proposition.

1. No new information/repetition of proposition. Elaborations coded into this category provide no further information over what is already provided in the proposition. The elaboration either restates the link or does not expand on the relationship presented in the link. For instance, for the proposition *muscle moves eye*, one student wrote "The muscles make it possible to move your eye."

2. Elaboration and proposition both correct. Elaborations in this category contain additional information that clarifies or elaborates on the correct relationship in the proposition. Explanations involve more detailed descriptions than in the link itself or report information about processes directly related to the relationship in the link. For example, the link *pupil is connected to iris* was explained as "The iris controls the amount of light that goes into the pupil."

3. Elaboration and proposition both incorrect. These elaborations contain incorrect additional information that adds to the incorrect proposition. The link *lens bends/refracts retina* was explained by one student as "This is the part that gets the message to the brain."

4. Correct elaboration paired with incorrect proposition. This category is particularly important because it identifies students who appear to understand the concept but have incorrectly constructed a proposition in their map. For instance, one student explained the proposition *brain sends messages to the optic nerve* as "The optic nerve is made to send all of the messages to the brain, that is why it is connected to the brain" (student incorrectly reversed direction of arrow in proposition).

5. Incorrect elaboration paired with correct proposition. This final category captures propositions that make sense paired with elaborations that add incorrect information. For example, *stirrup is a part of ear* was elaborated as "The stirrup moves when the eardrum vibrates."

Essay tasks. Administration instructions for the essay tasks were similar to the knowledge mapping instructions. Students were asked to write an essay explaining the most important ideas about the ear and the hearing process (or the eye and the vision process) that they would want an imaginary friend to understand. The essay

prompt also emphasized the relationships between the different parts of the ear (or eye) and how the ear (or eye) as a whole worked.

Essay tasks were scored using a holistic scoring rubric (and anchor papers) based on previous work (e.g., Baker, 1991, 1994; Baker, Aschbacher, et al., 1992; Baker, Freeman, & Clayton, 1991). Essays were scored on a scale of 1 to 5, with a score of 5 assigned to essays that covered the main scientific principles and included few or no misconceptions, and a score of 1 assigned to essays that showed little or no scientific understanding.

Four researchers (two pairs of two raters each) were involved in this process. Each pair of researchers worked together to select anchor papers and create scoring rubrics for one of the topic areas (hearing or vision). Then, each pair double-coded the essays in the other topic area (i.e., the pairs switched topic areas). Raters assigned each essay a score that ranged from 1 to 5, using the scoring rubric and sample anchor papers provided by the other rating team as guidelines. Thus, each essay task was assigned two scores, which we used to calculate interrater reliability figures. Reliability between raters was high for both essay tasks ($\alpha = .95$ for both hearing and vision).

Multiple-choice tasks. Students were instructed to complete the multiple-choice test using their knowledge about the hearing and vision processes. The test was split into two sections, and scored on a scale of 0 to 13 for each topic area.⁷ Inter-item alpha reliability was found to be .77 for hearing and .34 for vision.

Results

Descriptive Analyses

Table 4 shows the means and standard deviations for all tasks. Multiple-choice scores were high, with mean scores of 10.5 and 9.5 (out of a possible 13) on hearing and vision, respectively. Many students also performed well on the essays, with about 60% of the hearing essays and just over 40% of the vision essays scoring a 4 or 5 (on a 5-point scale). Essay means were 3.5 for hearing and 3.2 for vision. Hearing knowledge map scores ranged from 0 to 8.5, with a mean score of 2.7. Vision knowledge map scores ranged from 0 to 4, with a mean of 1.4.

⁷ Two of the test items (one hearing, one vision) asked students to order a set of steps in the hearing or vision process. Unlike the rest of the multiple-choice items, these order items were each worth 2 (rather than 1) point, thus making a total of 13 points per topic area.

Table 4

Means and Standard Deviations for Multiple-Choice, Essay, and Knowledge Mapping Tasks

	<i>n</i>	Maximum possible	<i>M</i>	<i>SD</i>
Multiple choice*				
Hearing	55	13	10.5	2.5
Vision	55	13	9.5	1.5
Essay**				
Hearing	55	5	3.5	1.0
Vision	54	5	3.2	1.0
Knowledge map*				
Hearing	53	14.5	2.7	2.2
Vision	52	13.5	1.4	1.1

* $p < .01$. ** $p < .05$.

Comparative Analyses

Results showed that the topic of vision was, in general, more difficult for the students than the topic of hearing. On both the multiple-choice and essay measures, students performed better on hearing than they did on vision: ($M = 10.5$ vs. $M = 9.5$), $t(54) = 11.9$, $p = .001$ for multiple-choice tasks; ($M = 3.5$ vs. $M = 3.2$), $t(53) = 6.5$, $p = .014$ for essay tasks. Knowledge mapping scores are in line with these trends: Students scored higher on hearing knowledge maps than they did on vision maps ($M = 2.7$ vs. $M = 1.4$, $t(49) = 20.6$, $p < .001$). Thus, the knowledge mapping task was sensitive to a difference in students' competency level. That is, results in knowledge mapping are consistent with other measures of topic understanding, providing evidence for the validity of the knowledge mapping measures.

MTMM Analyses

Our 3×2 design included two traits (understanding of the hearing process and understanding of the vision process) measured by three methods (multiple choice, essay, and knowledge map). The MTMM matrix is shown in Table 5. As Campbell and Fiske (1959) suggest, there are four basic properties to look for in the MTMM matrix: (a) high main diagonal reliabilities (e.g., reliabilities of each of our three tasks), (b) high correlations between different measures of each trait (e.g., between all measures of hearing), (c) low correlations between different measures of different traits (e.g., between multiple-choice, essay, and mapping tasks in different topic

Table 5

Multitrait-Multimethod Validity Matrix

		Multiple choice		Essay		Knowledge map	
		Hearing	Vision	Hearing	Vision	Hearing	Vision
Multiple choice	Hearing	.74	.49	.71	.66	.46	.37
	Vision	.49	.34	.61	.53	.48	.33
Essay	Hearing	.71	.61	.95	.59	.56	.47
	Vision	.66	.53	.59	.95	.39	.43
Knowledge map	Hearing	.46	.48	.56	.39	.72	.38
	Vision	.37	.33	.47	.43	.38	.77

areas), (d) smaller correlations between assessments measuring different traits (e.g., between mapping tasks in hearing and vision) than between those assessments measuring the same trait (e.g., essay and mapping tasks within hearing). We discuss each of these four properties in turn.

1. Main diagonal reliabilities. The main diagonal of the MTMM matrix shows the reliability of each measure. As described in the previous section, our reliability estimates were moderately high for knowledge mapping (.72 for hearing, .77 for vision) and essay tasks (.95 for both hearing and vision). Multiple-choice test reliability varied by topic area: Hearing items had an alpha reliability of .77, but vision items had a reliability of only .34. This low reliability on the vision multiple-choice test can perhaps be explained by a ceiling effect: Because students performed very well on almost all of the vision items, the variance was low, reducing the scale's reliability. With this exception, we meet the first MTMM criterion.

2. Correlations among different measures of each trait. To meet this criterion, we examine how well the three types of measures (multiple choice, essay, and knowledge map) agreed within each content area. Looking first at the measures of hearing knowledge, we see a correlation of .71 between the essay and multiple-choice tasks, a correlation of .46 between multiple-choice and knowledge mapping tasks, and a correlation of .56 between essay and mapping tasks. Thus, the knowledge mapping task correlates more highly with the essay than with the multiple-choice task. Note that the correlation between the essay and multiple-choice tasks is the highest—suggesting that our multiple-choice and essay tasks may have been measuring something quite similar.

Turning to the vision measures, we see correlations of .53, .33, and .43, respectively, between the essay and multiple-choice tasks, the multiple-choice and knowledge mapping tasks, and the essay and knowledge mapping tasks. Again, the mapping task correlates more highly with the essay than with the multiple-choice task. As we expect knowledge mapping to measure something more akin to essay knowledge than multiple-choice knowledge, these correlations may bolster the convergent validity of the essay and mapping measures. Again, the higher correlation between essay and multiple-choice tasks may indicate the similarity of those two measures.

3. Correlations among different measures of different traits. Here we examine how multiple-choice, essay, and knowledge-map measures correlate across the topics of hearing and vision. We would expect low correlations between different measures of these two different topics. Reading from the matrix, we have the following correlations in this category: .66, .61, .37, .48, .47, and .39. Another trait (e.g., students' general achievement, students' science ability) may explain why these correlations are moderate rather than low. That is, "good students" might understand both hearing and vision, explaining why we fail to meet this MTMM criterion for the traits of hearing and vision.

4. Between-trait vs. within-trait correlations. We expect that measures attempting to capture information about students' understandings *within* a particular topic area (monotrait) will correlate more highly than those measures of understanding *across domains* (heterotrait). Thus, we are comparing those monotrait matrix entries discussed in criterion 2 (.71, .46, .56, .53, .33, and .43) with correlations within each measure (multiple choice, essay, map). Reading from the matrix, we have the following heterotrait correlations between multiple-choice measures of hearing and vision, essay measures of hearing and vision, and knowledge mapping measures of hearing and vision, respectively: .49, .59, and .38. Our data violate this last MTMM assumption, with both sets of correlations being roughly equivalent. This may be due to the differences in what our monotrait measures assess: We clearly expect multiple-choice, essay, and knowledge mapping tasks to target slightly different aspects of students' understandings. Thus, equivalent sets of correlations may not be surprising.

Our MTMM results are inconclusive. We meet certain constraints, while possibly violating others. The lack of reliability of our vision multiple-choice task could account for some of the violations.

Additional Supporting Analyses

Two additional sets of analyses were conducted in order to further investigate the validity of knowledge mapping as an assessment tool. Students' propositions were compared to those propositions selected by experts as "critical links" to verify that students were indeed including significant links when constructing knowledge maps. In addition, students' explanations, or elaborations, of the links they considered most important were analyzed to better understand what information might be lost by using knowledge mapping rather than written explanations.

Proposition analysis. An additional measure of the validity of knowledge maps is the relevance and significance of the propositions created by students. Do students demonstrate their understandings by constructing "critical" links, as described by experts? To this end, we asked our content experts to identify the key or crucial relationships in their knowledge maps. We then examined the number of students who included each of these critical propositions in their maps. Table 6 shows these propositions and the percentage of students who included each in their hearing and vision maps.

Table 6

Percentage of Students Who Used Expert-Specified "Critical" Propositions for Hearing and Vision ($n = 53$ for hearing; $n = 52$ for vision)

Proposition	% of students who used this proposition
Hearing	
Nerve sends messages to brain	42
Sound waves vibrate eardrum	29
Eardrum vibrates hammer	29
Stirrup vibrates cochlea	4
Cochlea is connected to nerve	4
Fluid bends hair	0
Vision	
Optic nerve sends messages to brain	50
Light hits object	21
Light hits retina	8
Lens focuses light	2
Retina is connected to optic nerve	0
Brain creates image	0

In line with comparative analysis results, critical-proposition analysis results showed students were more proficient in hearing than vision. Five of the six critical propositions in hearing were used by students. The most-used hearing proposition, *nerve sends messages to brain*, was included in 42% of student maps; two other critical propositions were included in 28% of student maps, and the two least-used propositions appeared in only 4% of student maps. Students' vision knowledge maps were less likely to include expert-defined critical links. They used four out of six vision critical links. Of these links, the most popular was *optic nerve sends messages to brain*, included by 50% of the students in their vision maps. A second popular proposition was used by 21% of students, with two additional links used by 8% and 2% of the students, respectively.

We also examined the most popular links in student maps. What kinds of propositions were students likely to create and how "good" were these links? Of the five most popular hearing links created by students, only two were not critical links. Furthermore, these two noncritical popular hearing links were found in expert maps. In vision (the topic area in which students showed less understanding on all measures), three of the five most popular student-constructed links were not critical links. Of these three, only one was found in the experts' maps; the other two propositions characterized low-level, basic relationships not found in expert maps, but still correct. These data demonstrate that in general students were including important propositions in their knowledge maps—links found in expert maps and identified by experts as particularly significant to the hearing or vision process—with hearing the better understood process among students.

Elaboration analysis. Table 7 shows (by topic area) the percentage of student elaborations that fell into each of the five categories described in our method section. Most elaborations neither increased nor diminished our estimate (based solely on students' knowledge maps) of students' scientific understandings. Almost half of the students' elaborations (46% for hearing and 49% for vision) fell into the first category; that is, the explanations simply repeated the knowledge map propositions, giving us no further insight into student's level of scientific understanding. Like category 1, category 3 elaborations (10% to 11% of elaborations) gave us no additional information, as both the proposition and the explanation were incorrect. Examining category 5, we found that only about 2% of student elaborations for hearing and vision paired a correct link with an incorrect understanding of the

Table 7

Percentage of Student Elaborations in Hearing and Vision by Elaboration Category

Elaboration category	% of hearing elaborations ^a	% of vision elaborations ^b
1. No new information/ repetition of proposition	46	49
2. Elaboration and proposition both correct	34	27
Information elsewhere in map	13	6
Information not possible in map	15	10
Information not in map	7	11
3. Elaboration and proposition both incorrect	11	10
4. Correct elaboration paired with incorrect proposition	6	12
Information elsewhere in map	0	1
Arrow wrong direction	2	1
Information not possible in map	1	2
Compound sentence	1	2
Information not in map	1	5
5. Incorrect elaboration paired with correct proposition	2	2

Note. Percentages may not sum to 100 due to rounding errors.

^aTotal number of hearing elaborations = 80.

^bTotal number of vision elaborations = 92.

underlying concept—showing that knowledge mapping is not likely to overestimate students' knowledge.

However, category 2 and category 4 elaborations may indicate underestimation by knowledge mapping of students' knowledge. About one third of the student elaborations (34% for hearing and 27% for vision) clarified an already correct proposition. These category 2 elaborations involved greater understanding than was seen in their corresponding propositions and thus bear further investigation. Are knowledge maps underestimating students' knowledge?

In order to explore this possibility, we further categorized these category 2 elaborations into three subgroups: (a) elaborations demonstrating understanding shown elsewhere in the map (no underestimation), (b) elaborations demonstrating understanding not possible to convey in the map (possible underestimation), and (c) elaborations demonstrating understanding not shown elsewhere in the map (underestimation).

Some of the category 2 elaborations (13% of all hearing elaborations, 6% of all vision elaborations) contained information not found in the corresponding links but found elsewhere in the student maps. These elaborations therefore were not showing new information, and the maps were not underestimating students' understandings. Additional category 2 elaborations (15% of all hearing elaborations, 10% of all vision elaborations) contained information difficult or impossible to reflect in the knowledge map, given the constrained terms and links provided to students. One example of such an elaboration was "hair protects the ear because it keeps the dirt out," a fact experts did not consider important enough to include in their maps. Another example shows the difficulty some students had with proposition creation. The elaboration "the optic nerve sends messages to the brain to tell you what you have seen" was paired with the proposition *optic nerve sends messages to brain*. There are two ideas in this elaborating sentence. However, the second, missing proposition, *brain creates image*, may have been too abstract for children wanting to create a sentence such as "The brain recognizes it" or "The brain knows what you see." Likewise for hearing, many students included elaborations similar to one student's explanation of the proposition *nerve sends messages to brain*: "Because [the brain] determines what the sound is." Here, we see a conceptual understanding missed by this hearing knowledge map: The student was trying to convey that he or she understood the function of the brain in the hearing process—namely, to recognize the sound—however, experts did not include a term or link that would allow students to make a proposition such as *brain recognizes sound*. Thus, this subset of category 2 elaborations demonstrates a task-specific problem: A correctable error in task creation, rather than a general knowledge mapping problem, caused these category 2 elaborations to appear to underestimate students' knowledge. Task modification could solve these issues.

Finally, certain category 2 elaborations (7% of all hearing elaborations, 11% of all vision elaborations) conveyed understanding not found in the students' maps. Students including these elaborations understood more than they showed in their maps and were not given credit for these understandings, a clear underestimation of their knowledge.

Category 4 elaborations pair correct explanations with incorrect propositions. These elaborations bear further scrutiny as they too may indicate that the student knowledge maps underestimate students' actual knowledge. In 6% to 12% of all student explanations, students demonstrated that they understood a concept

although they incorrectly created a matched proposition. In one case (1%), the student's map reflected this understanding via another proposition. In addition, in 1% to 2% of cases, the incorrect link was related to the direction of its arrow, a problem we have found easy to correct with enhanced training. An additional 1% of hearing and 2% of vision elaborations indicated understanding that could not be shown in the knowledge map, due to lack of available concepts or links. None of these types of category 4 elaborations demonstrated new knowledge underestimated by the knowledge maps.

However, two additional subsets of category 4 elaborations do show underestimations of students' knowledge. In translating compound sentences (1% to 2% of student elaborations), students had trouble creating two matched links or deciding on the "object" of the sentence; thus, *sound waves* get lost in the link *ear canal vibrates eardrum*, elaborated as "the sound waves go through the ear canal and make the eardrum vibrate"—a correct understanding of an incorrectly constructed link (the propositions should have been *sound waves travel through ear canal* and *sound waves vibrate eardrum*). Finally, in 1% of hearing and 5% of vision elaborations, students used the elaboration to add valuable information unrelated to their incorrect link, showing a clear underestimation of students' knowledge.

In summary, analysis of students' elaborations showed that in the majority of cases, knowledge mapping seems to accurately assess students' scientific understanding, matching information in their written elaborations. However, in some cases knowledge mapping may be underestimating slightly this understanding. About 7% to 11% of student elaborations (falling into category 2) and 2% to 7% of student elaborations (falling into category 4) showed possible underestimation by the mapping task of students' scientific understandings.

Discussion and Conclusion

We have addressed three issues in order to establish the validity of knowledge mapping in an assessment setting. First, we have outlined a model for the development of new knowledge mapping assessment tasks in order to ensure the content validity of these measures. Next, we have suggested a scoring system that is valid on its face (i.e., comparison to experts seems reasonable), reliable (i.e., high expert agreement), and validated by other, more complicated scoring schemes. Lastly, we have conducted numerous analyses related to the knowledge mapping tasks themselves.

We found knowledge maps to be sensitive to differences in students' competency level, with students scoring higher on hearing tasks than on vision tasks. However, the results of our multitrait-multimethod validity analysis were, at best, mixed. The MTMM matrix main diagonal reliabilities were high with the exception of the vision multiple-choice task. Although the correlations were not high, the knowledge mapping task correlated more highly with the essay than with the multiple-choice tasks for both hearing and vision—lending support for the convergent validity of essay and knowledge mapping measures. The correlations between different measures of different traits were moderate rather than low; this might be explained by an achievement trait possibly underlying our data. We violated the last MTMM assumption—that correlations between assessments measuring different traits are smaller than those between assessments measuring the same trait—by having roughly equivalent correlations. However, this result is not surprising since the different measures assess different aspects of understanding—some more basic, some more conceptual. Thus, our MTMM analyses neither support nor refute our validity claim.

Analyses of the students' propositions and elaborations indicated that students understood what a link was, could explain their mapping components in writing, and—like experts—were constructing relevant and significant propositions in their knowledge maps. However, students' elaborations also indicated that knowledge maps might be slightly underestimating the level of scientific understanding of students.

Future work needs to address training issues with students. We believe that teaching students to select the “best” or deepest links between concepts (rather than any link that makes sense or is correct) will solve the problem of underestimating students' understanding of the domain. Further practice with knowledge mapping will also help students in two more basic areas: understanding the correct direction of the link arrows and splitting connected ideas into multiple propositions.

In conclusion, our results suggest that knowledge mapping is measuring something valuable, including some factual and some deeper conceptual scientific understandings. Further, this assessment task affords students the opportunity to show what they know in a new way. We believe we have presented a comprehensive model-based approach for the development and scoring of these knowledge mapping measures to foster validity. More research in this area will help

determine for whom knowledge mapping works best and how we might best utilize this assessment approach in the future.

References

- Armbruster, B. B., & Anderson, T. H. (1984). Mapping: Representing informative text diagrammatically. In C. D. Holley & D. F. Dansereau (Eds.), *Spatial learning strategies: Techniques, applications, and related issues* (pp. 189–209). Orlando, FL: Academic Press.
- Austin, L. B., & Shore, B. M. (1995). Using concept mapping for assessment in physics. *Physics Education*, 30, 41–45.
- Baker, E. L. (1991, April). *Expectations and evidence for alternative assessment*. Paper presented at the annual meeting of the American Educational Research Association, Chicago.
- Baker, E. L. (1994). Learning-based assessments of history understanding. *Educational Psychologist*, 29, 97–106.
- Baker, E. L., Aschbacher, P. R., Niemi, D., & Sato, E. (1992). *CRESST performance assessment models: Assessing content and explanation*. Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing.
- Baker, E. L., Freeman, M., & Clayton, S. (1991). Cognitive assessment of history for large-scale testing. In M. C. Wittrock & E. L. Baker (Eds.), *Testing and cognition*. Englewood, NJ: Prentice Hall.
- Baker, E. L., & Niemi, D. (1998). *Progress report 6: Pilot testing, scoring, and refinement of mathematics and language arts performance assessments* (Deliverable to LAUSD). Los Angeles: University of California, Center for the Study of Evaluation.
- Baker, E. L., Niemi, D., Novak, J., & Herl, H. (1992). Hypertext as a strategy for teaching and assessing knowledge representation. In S. Dijkstra, H. P. M. Krammer, & J. J. G. van Merriënboer (Eds.), *Instructional models in computer-based learning environments* (pp. 365–384). Berlin: Springer-Verlag.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81–105.
- Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, 4, 55–81.
- Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, 5, 121–125.
- Chi, M. T. H., Glaser, R., & Farr, M. J. (1988). *The nature of expertise*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Chung, G. K. W. K., O'Neil, H. F., Jr., & Herl, H. E. (1999). The use of computer-based collaborative knowledge mapping to measure team processes and team outcomes. *Computers in Human Behavior*, 15, 463–494.

- Gentner, D., & Stevens, A. L. (1983). *Mental models*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Glaser, R., & Bassok, M. (1989). Learning theory and the study of instruction. *Annual Review of Psychology*, 40, 631–666.
- Hardy, R. A. (1995). Examining the cost of performance assessment. *Applied Measurement in Education*, 8, 121–134.
- Heinze-Fry, J. A., & Novak, J. D. (1990). Concept mapping brings long-term movement toward meaningful learning. *Science Education*, 74, 461–472.
- Herl, H. E. (1995). *Construct validation of an approach to modeling cognitive structure of experts' and novices' U.S. history knowledge*. Unpublished doctoral dissertation, University of California, Los Angeles.
- Herl, H. E., Baker, E. L., & Niemi, D. (1996). Construct validation of an approach to modeling cognitive structure of U.S. history knowledge. *Journal of Educational Research*, 89, 206–219.
- Holley, C. D., & Dansereau, D. F. (1984). Networking: The technique and the empirical evidence. In C. D. Holley & D. F. Dansereau (Eds.), *Spatial learning strategies: Techniques, applications, and related issues* (pp. 81–108). Orlando, FL: Academic Press.
- Horton, P. B., McConney, A. A., Gallo, M., Woods, A. L., Senn, G. J., & Hamelin, D. (1993). An investigation of the effectiveness of concept mapping as an instructional tool. *Science Education*, 77, 95–111.
- Jonassen, D. H., Beissner, K., & Yacci, M. (1993). *Structural knowledge: Techniques for representing, conveying, and acquiring structural knowledge*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Markham, K. M., Mintzes, J. J., & Jones, M. G. (1994). The concept map as a research and evaluation tool: Further evidence of validity. *Journal of Research in Science Teaching*, 31, 91–101.
- McClure, J. R., Sonak, B., & Suen, H. K. (1999). Concept map assessment of classroom learning: Reliability, validity, and logistical practicality. *Journal of Research in Science Teaching*, 36, 475–492.
- Novak, J. D. (1998). *Learning, creating, and using knowledge: Concept maps as facilitative tools in schools and corporations*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Novak, J. D., & Gowin, D. B. (1984). *Learning how to learn*. New York: Cambridge University Press.
- Okebukola, P. A., & Jegede, O. (1988). Cognitive preference and learning mode as determinants of meaningful learning through concept mapping. *Science Education*, 72, 489–500.

- Osmundson, E. (1998). *Children as scientists: Adding, connecting, and elaborating ideas on the way to conceptual development*. Unpublished doctoral dissertation, University of California, Los Angeles.
- Osmundson, E., Chung, G. K. W. K., Herl, H. E., & Klein, D. C. D. (1999). *Concept mapping in the classroom: A tool for examining the development of students' conceptual understanding* (CSE Tech. Rep. No. 507). Los Angeles: University of California, Center for Research on Evaluation, Standards, and Student Testing.
- Ruiz-Primo, M. A., Schultz, S. E., & Shavelson, R. J. (1997). *Concept map-based assessment in science: Two exploratory studies* (CSE Tech. Rep. No. 436). Los Angeles: University of California, Center for Research on Evaluation, Standards, and Student Testing.
- Ruiz-Primo, M. A., & Shavelson, R. J. (1995, April). *Concept maps as potential alternative assessments in science*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.
- Ruiz-Primo, M. A., Shavelson, R. J., & Schultz, S. E. (1997). *On the validity of concept map-base assessment interpretations: An experiment testing the assumption of hierarchical concept maps in science* (CSE Tech. Rep. No. 455). Los Angeles: University of California, Center for Research on Evaluation, Standards, and Student Testing.
- Wallace, J. D., & Mintzes, J. J. (1990). The concept map as a research tool: Exploring conceptual change in biology. *Journal of Research in Science Teaching*, 27, 1033–1052.

Appendix A

Knowledge Mapping Task Prompts—Hearing and Vision

Name: _____

Class Time: _____

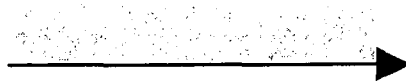
Hearing Knowledge Mapping Task

Your friend has missed the last two months of school and wants you to explain all about our ears and how they work. Using this mapping software, create a knowledge map to organize all of the important ideas about hearing and to show how the different parts of the hearing process go together with one another. Your knowledge map should include all of the information that your friend might need to know to really understand how our ears work.

1.



2.



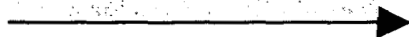
Name: _____

Class Time: _____

Vision Knowledge Mapping Task

Your friend has missed the last two months of school and wants you to explain all about our eyes and how they work. Using this mapping software, create a knowledge map to organize all of the important ideas about vision and to show how the different parts of the vision process go together with one another. Your knowledge map should include all of the information that your friend might need to know to really understand how our eyes work.

1.



2.



Appendix B

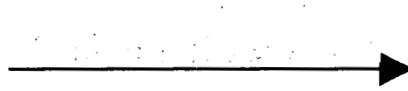
Elaboration Worksheet of Most Important Links

Name: _____

Class Time: _____

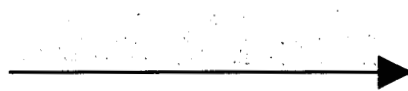
The two most important links in my map are:

1.



Explanation of link #1:

2.



Explanation of link #2:

Appendix C

Essay Task Prompts—Hearing and Vision

Name: _____

Class Time: _____

Hearing Essay Task

Imagine your friend comes to you with a problem. She has missed the last two months of school and wants you to explain how ears work. You need to explain all about the ear and the hearing process.

Think about all of the important things you've learned about hearing and how our ears work. Also think about the relationships between the different parts of the ear and how the ear as a whole goes together. Then write an explanation to your friend so that she can understand hearing.

- Write an essay explaining the most important ideas you want your friend to understand.
- Include what you've learned in class about the most important elements of hearing.
- Include both general concepts and specific facts that you know about hearing. Make sure you explain the purpose of the different parts of the ear so your friend will understand how the ear works.

After you have finished writing, you may want to reread your answer and make corrections.

Begin your essay on the next page.

Name: _____

Class Time: _____

Vision Essay Task

Imagine your friend comes to you with a problem. She has missed the last two months of school and wants you to explain how eyes work. You need to explain all about the eye and the vision process.

Think about all of the important things you've learned about vision and how our eyes work. Also think about the relationships between the different parts of the eye and how the eye as a whole goes together. Then write an explanation to your friend so that she can understand vision.

- Write an essay explaining the most important ideas you want your friend to understand.
- Include what you've learned in class about the most important elements of vision.
- Include both general concepts and specific facts that you know about vision. Make sure you explain the purpose of the different parts of the eye so your friend will understand how the eye works.

After you have finished writing, you may want to reread your answer and make corrections.

Begin your essay on the next page.

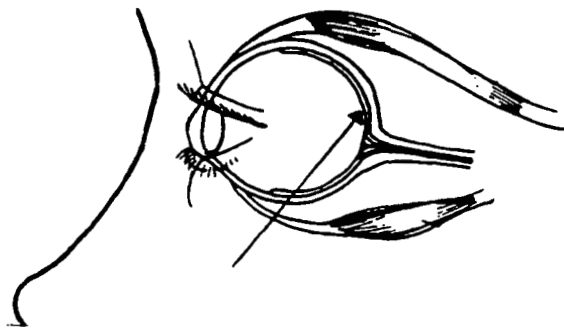
Appendix D
Multiple-Choice Test

Name: _____

Class Time: _____

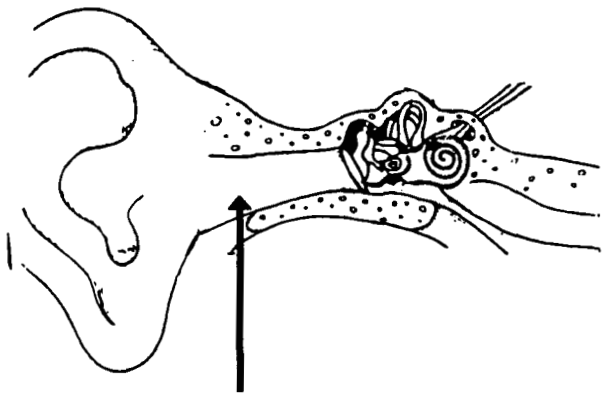
Hearing and Vision Multiple-Choice Task

1. Which part of your eye is really a hole?
 - a. iris
 - b. pupil
 - c. cornea
 - d. sclera
2. Which part of your eye helps you see color?
 - a. cornea
 - b. rods
 - c. cones
 - d. optic nerve
3. On the diagram of an eye, the part labeled with an arrow is the:
 - a. cornea.
 - b. iris.
 - c. pupil.
 - d. retina.
4. What is the round, colored part of your eye called?
 - a. iris
 - b. cornea
 - c. retina
 - d. lens
5. The purpose of the pupil is to:
 - a. control the amount of light.
 - b. focus the image.
 - c. send a message to the brain.
 - d. all of the above.
6. The optic nerve connects which two parts of the body?
 - a. the eyes and the ears
 - b. the ears and the brain
 - c. the ears and the nose
 - d. the eyes and the brain

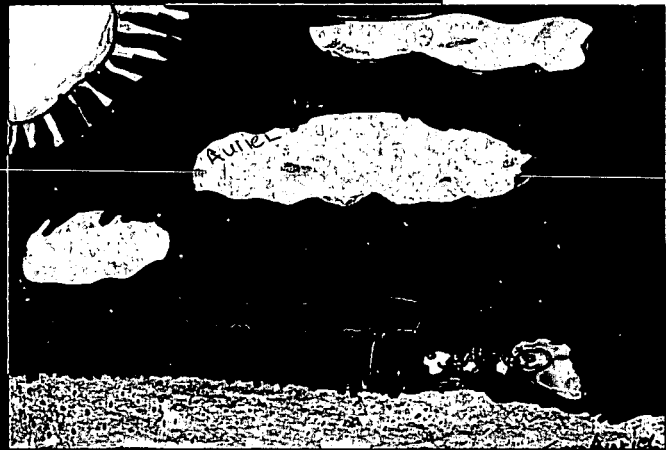
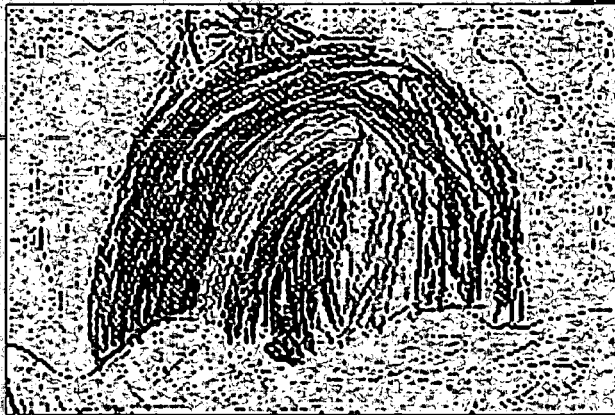
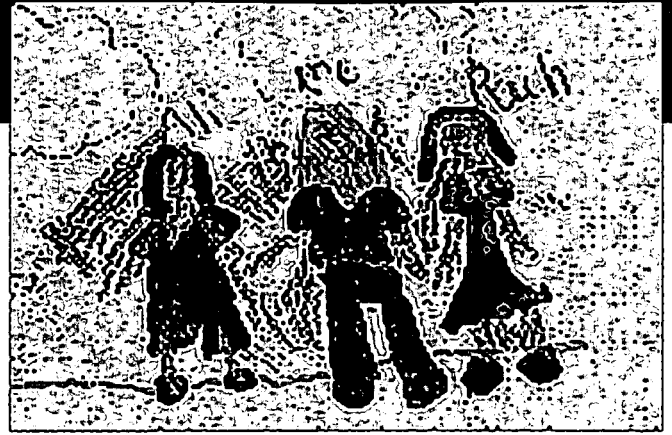


7. Which part of the eye bends the light into the eyeball?
 - a. iris
 - b. lens
 - c. retina
 - d. all of the above
8. *Reflection* is:
 - a. what you see in the mirror.
 - b. how light hits objects.
 - c. the way you see color.
 - d. all of the above.
9. If you walk out of the movies into bright light, what happens to your pupil?
 - a. It vibrates.
 - b. It gets smaller.
 - c. It disappears.
 - d. All of the above
10. Which of the following are the muscles in the eye not responsible for?
 - a. Moving the eye
 - b. Opening and closing the pupil
 - c. Turning the image right side up
 - d. Helping the lens to focus light
11. Imagine you see a tree outside. How do you know it's a tree?
 - a. The image is focused on the retina, which decides it's a tree.
 - b. The image is focused on the retina, then sent to the brain, which decides it's a tree.
 - c. The image is focused on the lens, then sent to the retina, which decides it's a tree.
 - d. The image is focused on the retina, then sent to the optic nerve, which decides it's a tree.
12. Put the following statements into the correct order by writing the appropriate numbers in the spaces provided. Number from 1 to 5.
___ Light passes through the cornea.
___ A message is sent to the brain.
___ Light passes through the pupil.
___ Light is bent into the eyeball.
___ Light is focused on the retina.

13. Which part of the ear is responsible for catching, or collecting, the sound waves?
- outer ear
 - middle ear
 - inner ear
 - ear drum
14. Which part of the ear is filled with fluid?
- hammer
 - outer ear
 - cochlea
 - all of the above
15. How does an eardrum "hear?"
- It feels beeps.
 - It sees waves.
 - It senses patterns.
 - It feels vibrations.
16. Sound waves travel through:
- air.
 - liquid.
 - plastic.
 - all of the above.
17. On the diagram of an ear, the part labeled with an arrow is the:
- cochlea.
 - stirrup.
 - hammer.
 - ear canal.
18. When there is sound, which three bones vibrate in the ear?
- ear canal, hammer, and cochlea
 - hammer, stirrup, and anvil
 - stirrup, inner ear, and anvil
 - cochlea, stirrup, and eardrum



19. Which of the following is not part of the middle ear?
- stirrup
 - hammer
 - cochlea
 - anvil
20. How are messages about what sounds we hear carried to the brain?
- ear canal
 - auditory nerve
 - inner ear
 - anvil
21. Which part of the ear is made up of a thin membrane?
- eardrum
 - anvil
 - stirrup
 - middle ear
22. Which part of the ear looks like a spiral?
- eardrum
 - auditory nerve
 - cochlea
 - ear canal
23. Put the following statements into the correct order by writing the appropriate numbers in the spaces provided. Number from 1 to 5.
- ___ Sound waves travel through the ear canal.
- ___ Message is sent to the brain.
- ___ Sound waves cause the eardrum to vibrate.
- ___ Sound waves enter the outer ear.
- ___ Bones in the middle ear continue vibrations.
24. Which part of the ear helps you keep your balance?
- outer ear
 - middle ear
 - inner ear
 - all of the above





U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)



TM034190

NOTICE

Reproduction Basis



This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.



This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").